Received: 5 June 2017 | Accepted: 5 June 2017

DOI: 10.1111/jep.12792

WILEY Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

ORIGINAL ARTICLE

Minimizing imbalances on patient characteristics between treatment groups in randomized trials using classification tree analysis

Ariel Linden DrPH[1,2] | Paul R. Yarnold PhD[3]

[1] President, Linden Consulting Group, LLC, Ann Arbor, Michigan, USA

[2] Research Scientist, Division of General Medicine, Medical School−−University of Michigan, Ann Arbor, Michigan, USA

[3] President, Optimal Data Analysis, LLC, Chicago, Illinois, USA

Correspondence
Ariel Linden, Linden Consulting Group, LLC, 1301 North Bay Drive, Ann Arbor, MI 48103, USA.
Email: alinden@lindenconsulting.org

Abstract

Rationale, aims, and objectives: Randomization ensures that treatment groups do not differ systematically in their characteristics, thereby reducing threats to validity that may otherwise explain differences in outcomes. Large observed imbalances in patient characteristics may indicate that selection bias is being introduced into the treatment allocation process. We introduce classification tree analysis (CTA) as a novel algorithmic approach for identifying potential imbalances in characteristics and their interactions when provisionally assigning each new participant to one or the other treatment group. The participant is then permanently assigned to the treatment group that elicits either no or less imbalance than if assigned to the alternate group.

Method: Using data on participant characteristics from a clinical trial, we compare 3 different treatment allocation approaches: permuted block randomization (the original allocation method), minimization, and CTA. Treatment allocation performance is assessed by examining balance of all 17 patient characteristics between study groups for each of the allocation techniques.

Results: While all 3 treatment allocation techniques achieved excellent balance on main effect variables, Classification tree analysis further identified imbalances on interactions and in the distributions of some of the continuous variables.

Conclusions: Classification tree analysis offers an algorithmic procedure that may be used with any randomization methodology to identify and then minimize linear, nonlinear, and interactive effects that induce covariate imbalance between groups. Investigators should consider using the CTA approach as a real-time complement to randomization for any clinical trial to safeguard the treatment allocation process against bias.

KEYWORDS

classification tree analysis, clinical trials, machine learning, randomization

1 | INTRODUCTION

The randomized controlled trial (RCT) is considered the gold standard study design because randomization ensures that treatment groups do not differ systematically in their characteristics. However, the randomization process may still produce between-group differences that may explain observed differences in outcomes and thus constitute a threat to statistical conclusion validity.[1,2]

A common approach for assessing the effectiveness of the randomization process is to examine the balance (ie, comparability between distributions) of observed baseline characteristics among study groups.[3,4] Whereas a few small imbalances are expected by chance, larger imbalances may be indicative of selection bias in the treatment allocation process. The latter case necessitates the use of techniques that model the treatment assignment to adjust for these imbalanced characteristics that otherwise compromise causal interpretation of the results (see for example, previous studies[5-9]).

Given the large investment and organizational complexity required to properly conduct an RCT, it is prudent for study administrators to use strategies throughout the entire trial that safeguard the treatment

allocation process against systematic bias. *Minimization* is a strategy specifically designed to reduce imbalances between treatment groups on a predefined set of categorized patient characteristics.[10,11] It is implemented after the first participant enrolled in the study is randomly assigned to a treatment group and allocates each subsequent participant to the treatment arm that minimizes the imbalance on one or more of the selected characteristics, cumulatively until that point in time. The procedure allows for probabilistic treatment allocation (where 1.0 means that the patient is always assigned to the group that minimizes imbalances, and 0.50 is equivalent to a coin flip) as well as the ability to assign weights to characteristics deemed more important than others (eg, variables thought to be prognostic of the outcome).[11]

In a systematic review of the minimization literature, Scott et al[12] report that in most cases, minimization outperforms simple randomization in achieving balanced groups (particularly when trial sample sizes are small), and it holds an advantage over stratified randomization methods due to its ability to incorporate more prognostic factors. Moher et al[3] contend that minimization offers the only acceptable alternative to randomization, while Treasure and MacRae[13] assert that minimization is actually superior to randomization.

In this paper, we introduce classification tree analysis (CTA)[14,15] as a novel and more robust approach to minimizing imbalances. Classification tree analysis is a machine learning algorithm that can identify potentially complex patterns in the data that distinguish patients assigned to the different treatment arms in a clinical trial. In contrast to existing techniques that require the investigator to determine which variables to include, how to categorize continuous and multicategorical variables, and which variables should be interacted (if any), CTA performs all these functions algorithmically, and if/as necessary, revises the allocation as each new participant is enrolled in the trial to minimize imbalances.

This paper is organized as follows. In Section 2, we describe the data used in the current study, provide a brief introduction to CTA, and explain the analytic framework used to minimize imbalances on patient characteristics. Section 3 reports and compares the results of the proposed CTA approach (which is implemented as a complement to randomization) to that of the original block randomization and to the commonly used minimization algorithm proposed by Pocock and Simon.[11] Section 4 describes the specific advantages of the CTA approach over existing approaches and considers potential limitations.

2 | METHODS

2.1 | Data

We use data from a parallel-group, stratified, clinical trial that examined whether a comprehensive, hospital-based, transitional care intervention reduces readmissions for participants with congestive heart failure (CHF) and chronic obstructive pulmonary disease.[16] The intervention involved nurses implementing motivational interviewing-based health coaching to improve patients' health behaviours, which in turn was expected to empower patients to better manage their own healthcare and reduce unplanned readmissions.[17-19] For the present study, we limit our analysis to the CHF cohort (N = 257), for which baseline characteristics include demographic variables (gender, age, insurance type, and living conditions); patient activation measure (PAM) score; prior year hospital utilization (admissions, average length of stay, and emergency department visits) both for CHF as well as for all causes; average length of stay for the index hospitalization; and the presence of 7 key comorbidities (chronic obstructive pulmonary disease, cerebrovascular disease, acute myocardial infarction, diabetes, renal disease, chronic pain, and obesity).

2.2 | Brief introduction to CTA

In its simplest form, CTA is an optimal discriminant analysis (ODA) model.[20] For any given application, the ODA algorithm identifies the cutpoint for an ordered attribute (independent variable), or the assignment rule for a categorical attribute, that most accurately (optimally) discriminates between 2 (or more) categories of the class (dependent variable).[21] This entails computing the effect strength for sensitivity (ESS) obtained using every possible cutpoint along the continuum of values (or every possible rule) to classify sample observations. Effect strength for sensitivity is the mean sensitivity obtained over class categories at the cutpoint (or via the rule) used to classify observations, standardized to a 0% to 100% scale on which 0% represents the level of accuracy expected by chance; 100% represents perfect accuracy; and negative values indicate accuracy worse than expected by chance. By definition, the optimal model uses the assignment rule that yields the greatest ESS value. Statistical significance is assessed via permutation probability, and validity analyses (eg, jackknife and holdout) are conducted to estimate potential cross-generalizability of the model in correctly classifying new subjects that may differ in their characteristics compared to subjects in the original sample.[14,22] Classification tree analysis constructs optimal models by chaining multiple ODA models together.[23] Classification tree analysis models classify observations into 1 of 2 or more subgroups represented as model endpoints (terminal nodes) called "strata" because the sample is stratified into subgroups that are homogeneous within, and heterogeneous between, the different endpoints defined by the attributes and corresponding optimal cutpoints/rules selected by the CTA model.[14,24]

2.3 | Analytic approach

This study compares the efficacy of 3 different treatment allocation approaches for producing treatment cohorts that are balanced on observed baseline characteristics.

2.3.1 | Permuted block randomization

The first treatment allocation approach is the original randomization process used in Linden and Butterworth[16] and serves as the basis for comparison. Prior to study commencement, a randomization sequence was generated to allocate participants to treatment arms using random permuted blocks.[25] There were 4 strata (2 hospitals and 2 disease conditions) and 5 permuted blocks allocated in equal proportions, with a minimum size of 2 and maximum size of 10. The treatments were allocated in a 1:1 ratio, with 18 extra allocations provided to maintain the integrity of the final block in each stratum. The allocation sequence was concealed via sequentially numbered, opaque sealed envelopes.

After each participant signed the consent form and provided baseline information, the envelope was opened by study staff in their presence, simultaneously revealing the treatment allocation to both participant and study staff.

2.3.2 | Minimization

The second treatment allocation approach applies the Pocock and Simon[11] minimization technique, which requires all variables ("factors") to be categorized. Given that there is no rule of thumb regarding how best to categorize continuous variables, we generated categories with relatively few levels that maintain sufficient sample size. The 7 continuous variables were categorized as follows. Age was categorized into 4 age-bands: 20 to 39; 40 to 59; 60 to 79; and 80+. PAM scores were categorized into 3 quantiles: 16.5 to 47.4; 47.5 to 56.4; and 56.5 to 100. All-cause hospitalizations were categorized into 4 levels: 1, 2, 3, and 4+. All-cause emergency department visits were categorized into 5 levels: 0, 1, 2, 3, and 4+. Congestive heart failure–specific hospitalizations were categorized into 4 levels: 0, 1, 2, and 3. Congestive heart failure–specific length of stay was categorized into 4 levels: 0; 1 to 3; 4 to 6; 7+. And finally, the length of stay of the index admission was categorized into 3 levels: 1 to 3; 4 to 6; and 7+. All categorical variables were left in their original number of levels.

Each study participant was assigned to treatment in the same sequence in which they were originally enrolled in the trial. Each of the 257 participant's 17 factors were entered into the Stata user-written command RCT_MINIM,[26] with the first participant randomly assigned to treatment, and all subsequent participants automatically allocated to the treatment arm that produced the least imbalance between groups. No weights were used in the computations, given the extensive number of variables available and the belief that they all have prognostic value.

2.3.3 | CTA-based minimization

The original study randomization sequence (see Section 2.3.1) served as the basis for the CTA approach. This ensures that the treatment assignment process cannot be deciphered (and the allocation process potentially compromised), even if CTA indicates that changes to individual allocations must be made along the way. The first 6 participants were randomized according to the original sequence, because a minimum sample of 7 is needed to identify a statistically significant model.[14]

Commencing with the seventh participant, a CTA model was sought in which the 17 independent variables (in their original measurement scale) were used to discriminate between treatment groups––including all 6 participants enrolled in the trial thus far, and including the seventh participant allocated to the treatment arm by block randomization sequencing. If no CTA model could be generated (ie, indicating that no imbalances were found across any variable or interactions between subsets of variables), then the participant was assigned to the treatment group as was originally intended. However, if a CTA model could be generated, then the participant was provisionally given the alternate treatment assignment and a CTA model was again sought. If no CTA model could be generated under this provisional assignment, then the participant would be permanently allocated to this treatment arm. In the case where a CTA model could be generated for both provisional treatment assignments, then the participant would be permanently allocated to the arm that elicited the CTA model with the lowest ESS value (indicating that this treatment assignment was less biased than the alternate assignment). This process was repeated sequentially for each of remaining individuals in the study.

To maximize expositive clarity, CTA models generated by the process are provided to illustrate interactive and parabolic imbalances that were identified, and corrected, by the CTA approach.

2.4 | Performance metrics for treatment allocation approaches

To assess the performance of the 3 treatment allocation approaches, we examine how well each one avoided imbalances in baseline characteristics between treatment groups for the total study population, using 2 different methods.

First, we apply the conventional method of testing for differences, using χ^2 tests for categorical variables and t tests for continuous variables.

Second, we apply the method described by Linden and Yarnold[27] that uses ODA to assess balance of characteristics between groups. In contrast to conventional tests for difference in means (or proportions), the ODA approach is insensitive to skewed data and outliers, and it additionally identifies a cutpoint (or rule) along the distribution of the characteristic that distinguishes between treatment assignments. The underlying assumption is that if treatment groups cannot be distinguished based on the distribution of each characteristic, then the treatment allocation was successful. In this framework, sensitivity, specificity, and ESS are used as balance diagnostics.

3 | RESULTS

Table 1 presents the baseline characteristics of the study population, under all 3 treatment allocation methods. As shown, there is virtually no difference among methods in achieving close balance between treatment groups using conventional statistical tests.

Table 2 presents the baseline characteristics of the study population under all 3 treatment allocation methods, analysed using ODA. Summary values represent the cutpoint (or rule) on the characteristic. Sensitivity is presented for the intervention group, and specificity is presented for the group assigned to usual care (control condition). All 3 allocation methods achieved balance for all of the characteristics under study (as indicated by a consistently weak ESS across all characteristics, and supported by nonstatistically significant P values >.05). With the exception of age, ODA identified the same cutpoint for all characteristics in the CTA approach and in the original block randomization (with similar sensitivity/specificity values), while identifying different cutpoints (and sensitivity/specificity values) for the minimization technique. In all, CTA identified 8 cases of imbalance, all of which were successfully eliminated in the following (subsequent) step of the procedure.

TABLE 1 Baseline characteristics of study participants, assigned to treatment under 3 different allocation methods

Characteristic	Block randomization			Minimization			Classification tree analysis		
	Treatment (N = 129)	Control (N = 128)	*P*	Treatment (N = 129)	Control (N = 128)	*P*	Treatment (N = 125)	Control (N = 132)	*P*
Female	64 (49.6)	66 (51.6)	.75	64 (49.6)	66 (51.6)	.75	62 (49.6)	68 (51.5)	.76
Age, mean (SD)	67.16 (13.38)	69.55 (12.64)	.14	68.49 (13.03)	68.22 (13.08)	.87	67.30 (13.12)	69.35 (12.91)	.21
Insurance									
1. Medicare	92 (71.3)	96 (75.0)	.69	93 (72.1)	95 (74.2)	.70	88 (70.4)	100 (75.8)	.35
2. Medicaid	10 (7.8)	7 (5.5)		7 (5.4)	10 (7.8)		10 (8.0)	7 (5.3)	
3. Commercial	14 (10.9)	16 (12.0)		16 (12.4)	14 (10.9)		13 (10.4)	17 (12.9)	
4. None	13 (10.1)	9 (7.0)		13 (10.1)	9 (7.0)		14 (11.2)	8 (6.1)	
Living conditions									
1.With spouse/caregiver	88 (68.2)	88 (68.8)	.69	87 (67.4)	89 (69.5)	.71	83 (66.4)	93 (70.5)	.56
2. Alone	40 (31.0)	37 (28.9)		39 (30.2)	38 (29.7)		41 (32.8)	36 (27.3)	
3. Other	1 (0.8)	2 (1.6)		2 (1.6)	1 (0.8)		1 (0.8)	2 (1.5)	
4. Homeless	0 (0.0)	1 (0.8)		1 (0.8)	0 (0.0)		0 (0.0)	1 (0.8)	
PAM, mean (SD)	54.73 (14.93)	53.83 (12.87)	.60	53.17 (13.44)	55.41 (14.36)	.20	54.99 (15.00)	53.61 (12.85)	.43
Admissions, mean (SD)	1.87 (1.44)	1.71 (1.29)	.36	1.81 (1.44)	1.77 (1.29)	.78	1.86 (1.44)	1.73 (1.30)	.45
ED visits, mean (SD)	0.84 (1.86)	0.70 (1.60)	.54	0.82 (1.87)	0.72 (1.58)	.63	0.80 (1.77)	0.74 (1.70)	.79
CHF admissions, mean (SD)	0.46 (0.65)	0.40 (0.59)	.45	0.46 (0.67)	0.40 (0.57)	.45	0.46 (0.65)	0.40 (0.59)	.48
CHF hospital days, mean (SD)	2.02 (3.69)	2.04 (4.00)	.97	1.91 (3.19)	2.15 (4.41)	.63	2.02 (3.73)	2.04 (3.96)	.98
LOS index, mean (SD)	5.33 (4.70)	5.14 (3.80)	.72	5.32 (4.23)	5.16 (4.32)	.76	5.24 (4.67)	5.23 (3.87)	.99
COPD	49 (38.0)	42 (32.8)	.39	46 (35.7)	45 (35.2)	.93	48 (38.4)	43 (32.6)	.33
CEVD	64 (49.6)	53 (41.4)	.19	58 (45.0)	59 (46.1)	.86	62 (49.6)	55 (41.7)	.20
Chronic pain	29 (22.5)	23 (18.0)	.37	28 (21.7)	24 (18.8)	.56	28 (22.4)	24 (18.2)	.40
Diabetes	64 (49.6)	54 (42.2)	.23	59 (45.7)	59 (46.1)	.95	61 (48.8)	57 (43.2)	.37
AMI	33 (25.6)	35 (27.3)	.75	33 (25.6)	35 (27.3)	.75	32 (25.6)	36 (27.3)	.76
Renal disease	44 (34.1)	32 (25.0)	.11	38 (29.5)	38 (29.7)	.97	42 (33.6)	34 (25.8)	.17
Obesity	54 (41.9)	44 (34.4)	.22	50 (38.8)	48 (37.5)	.84	52 (41.6)	46 (34.8)	.27

Abbreviations: AMI, acute myocardial infarction; CEVD, cerebrovascular disease; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; ED, Emergency Department; LOS, length of stay; PAM, patient activation measure. Values reported as no. (%) unless otherwise noted.

Figure 1 illustrates 4 cases in which CTA identified imbalanced characteristics when provisionally assigning a new patient to the prespecified (via the original permuted block randomization) treatment arm. As shown, these imbalances were due to interactions between variables or a parabolic function of a single variable. In Figure 1A, an interaction was identified between a participant's living condition and gender, while Figure 1B illustrates an interaction between a participant's number of prior year hospital days for CHF and the length of stay of the index hospitalization. Figure 1C,D indicates that age has a nonlinear relationship with treatment assignment.

4 | DISCUSSION

Although CTA has been recently applied to observational study data for improving causal inference,[24,27-33] this paper focuses on its use in support of participant assignment in clinical trials. Using data from a permuted block randomized RCT,[16] on the basis of conventional criteria, we found that both minimization and CTA methods achieved balance on observed characteristics equally as well as the randomization method used in the original study.

However, the CTA approach offers several key advantages over the other 2 methods that are not readily apparent from the baseline characteristics tables. First, CTA inherently finds interactions and nonlinear (eg, parabolic) effects among variables that are unlikely to be identified manually (Figure 1). As a result, CTA ensures that balance is achieved not only in main effects but also in all possible nonlinear and interactive effects––which could also serve to limit statistical conclusion validity.

Second, whereas minimization requires the investigator to create categories for continuous variables (or rules for categorical variables), the resulting cutpoints (assignment rules) are unlikely to yield the maximum possible discrimination between groups. In contrast, CTA intrinsically identifies the maximally accurate cutpoint (rule) for each variable, ensuring that variables are optimally categorized to maximize accurate discrimination of participants in different groups.[14]

Third, as new participants are enrolled in the study, CTA attempts to identify new models based on updated variable cutpoints (or rules) and new interactions. Thus, CTA is a dynamic assignment process, whereas minimization relies strictly on the factors as they were originally conceived. Finally, the CTA approach integrates with any randomization sequencing procedure (eg, the permuted block randomization process used in the present study), which mitigates concerns that treatment assignment derived using CTA may be predicted with certainty and thus potentially manipulated (which in turn, will impact the validity of study outcomes). This

TABLE 2 Baseline characteristics of study participants, assigned to treatment under 3 different allocation methods, evaluated using optimal discriminant analysis (ODA). Values represent cutpoints (or assignment rules for categorical measures) on the characteristic, and values in parentheses represent sensitivity (for treatment) and specificity (for controls), as a percent

	Block randomization				Minimization				Classification tree analysis			
Characteristic	Treatment (N = 129)	Control (N = 128)	ESS	*P*<	Treatment (N = 129)	Control (N = 128)	ESS	*P*<	Treatment (N = 125)	Control (N = 132)	ESS	*P*<
Age	≤74.6 (73.64)	>74.6 (38.28)	11.92	.30	>64.4 (69.77)	≤64.4 (36.72)	6.49	.94	≤59.1 (27.20)	>59.1 (84.09)	11.29	.36
Female	=1 (50.39)	=0 (51.56)	1.95	.81	=0 (50.39)	=1 (51.56)	1.95	.81	=1 (50.40)	=0 (51.52)	1.92	.80
Insurance	=2,4 (17.83)	=1,3 (87.50)	5.33	.62	=3,4 (22.48)	=1,2 (82.03)	4.51	.74	=2,4 (19.20)	=1,3 (88.64)	7.84	.30
Living cond.	=2 (31.01)	=1,3,4 (71.09)	2.1	.86	>1 (32.56)	=1 (69.53)	2.09	.88	=2 (32.80)	=1,3,4 (72.73)	5.53	.44
PAM	>58.2 (34.88)	≤58.2 (75.00)	9.88	.33	≤51.4 (54.26)	>51.4 (54.69)	8.95	.43	>58.2 (36.00)	≤58.2 (75.76)	11.76	.19
Admissions	>2 (18.60)	≤2 (85.94)	4.54	.71	>5 (3.88)	≤5 (97.66)	1.53	.99	>2 (18.40)	≤2 (85.61)	4.01	.77
ED visits	>1 (17.83)	≤1 (86.72)	4.55	.69	>1 (17.05)	≤1 (85.94)	2.99	.92	>2 (17.60)	≤2 (86.36)	3.96	.75
CHF admits	>0 (38.76)	=0 (65.62)	4.38	.59	>1 (6.98)	≤1 (96.09)	3.07	.77	>0 (38.40)	=0 (65.15)	3.55	.67
CHF days	>1 (38.76)	≤1 (67.97)	6.73	.49	>1 (37.21)	≤1 (66.41)	3.62	.93	>1 (38.40)	≤1 (67.42)	5.82	.62
LOS index	≤5 (69.77)	>5 (35.16)	4.92	.91	>5 (37.21)	≤5 (71.87)	9.08	.36	≤5 (70.40)	>5 (35.61)	6.01	.77
COPD	=1 (37.98)	=0 (67.19)	5.17	.44	=1 (35.66)	=0 (64.84)	0.5	.99	=1 (38.40)	=0 (67.42)	5.82	.37
CEVD	=1 (49.61)	=0 (58.59)	8.21	.21	=0 (55.04)	=1 (46.09)	1.13	.9	=1 (49.60)	=0 (58.33)	7.93	.22
Chronic pain	=1 (22.48)	=0 (82.03)	4.51	.44	=1 (21.71)	=0 (81.25)	2.96	.65	=1 (22.40)	=0 (81.82)	4.22	.45
Diabetes	=1 (49.61)	=0 (57.81)	7.42	.27	=0 (54.26)	=1 (46.09)	0.36	.99	=1 (48.80)	=0 (56.82)	5.62	.38
AMI	=0 (74.42)	=1 (27.34)	1.76	.78	=0 (74.42)	=1 (27.34)	1.76	.78	=0 (74.40)	=1 (27.27)	1.67	.78
Renal	=1 (34.11)	=0 (65.62)	9.11	.14	=0 (70.54)	=1 (29.69)	0.23	.99	=1 (33.60)	=0 (74.24)	7.84	.18
Obesity	=1 (41.86)	=0 (65.62)	7.49	.25	=1 (38.76)	=0 (62.50)	1.26	.91	=1 (41.60)	=0 (65.15)	6.75	.31

Abbreviations: AMI, acute myocardial infarction; CEVD, cerebrovascular disease; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; ED, Emergency Department; ESS, effect strength for sensitivity; LOS, length of stay; PAM, patient activation measure.

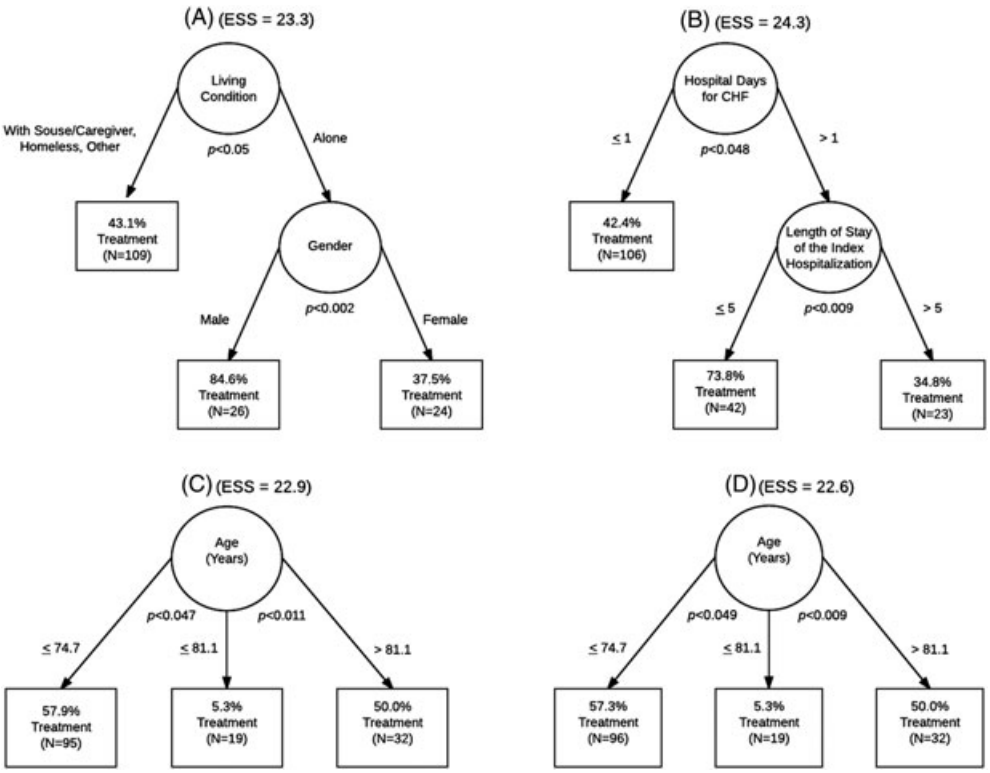

FIGURE 1 Interactive (A and B) and parabolic (C and D) classification tree analysis models identified for imbalanced subsamples. CHF, congestive heart failure; ESS, effect strength for sensitivity

advantage is verified by reviewing Table 2, which shows nearly identical values (for all metrics) were derived for the CTA and the block randomization approaches.

While the CTA approach has clear advantages, it also shares some of the same potential limitations as the minimization technique when compared to conventional randomization methods. First,

implementation of the procedure is naturally more complex than simply generating the randomization sequence prior to initiation of the study. Classification tree analysis requires real-time, manual entry of each new participant's characteristics into the software to determine treatment assignment. While this process is mechanically straightforward, it does necessitate that an administrator is available when participants are enrolled into the study. Another potential limitation is that a perfectly nonconfounded allocation sequence may not be possible. In this case, the outcome model should be adjusted for the nonrandom covariates (and interactions) identified in the allocation procedure to obtain accurate results. This concern has been raised in the context of the minimization technique,[12] and the use of CTA-based propensity scores to make such adjustments was recently demonstrated.[32]

5 | CONCLUSION

In summary, this paper introduces a novel machine learning approach for minimizing imbalances on patient characteristics between treatment groups in randomized trials. This approach offers several advantages over existing approaches, such as an algorithmic procedure to identify variables and interactions that induce imbalance, identifying the optimal (maximum accuracy) cutpoints (or classification rules) on those variables, and integrating with any randomization procedure to ensure validity of the study outcomes. Therefore, investigators should consider using the CTA approach as a real-time complement to randomization for any clinical trial as a means of safeguarding the treatment allocation process against selection bias.

ACKNOWLEDGEMENT

We wish to thank Julia Adler-Milstein for her review and suggestions on the manuscript.

REFERENCES

1. Shadish SR, Cook TD, Campbell DT. *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin; 2002.
2. Linden A. Estimating the effect of regression to the mean in health management programs. *Dis Manag Health Out*. 2007;15:7-12.
3. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
4. Linden A, Roberts N. A user's guide to the disease management literature: recommendations for reporting and assessing program outcomes. *American Journal of Managed Care*. 2005;11:81-90.
5. Linden A, Adams J. Evaluating disease management program effectiveness: an introduction to instrumental variables. *J Eval Clin Pract*. 2006;12:148-154.
6. Linden A, Adams JL. Using propensity score-based weighting in the evaluation of health management programme effectiveness. *J Eval Clin Pract*. 2010;16:175-179.
7. Linden A, Adams JL. Evaluating health management programmes over time. application of propensity score-based weighting to longitudinal data. *J Eval Clin Pract*. 2010;16:180-185.
8. Linden A. Combining propensity score-based stratification and weighting to improve causal inference in the evaluation of health care interventions. *J Eval Clin Pract*. 2014;20:1065-1071.
9. Linden A, Uysal SD, Ryan A, Adams JL. Estimating causal effects for multivalued treatments: a comparison of approaches. *Stat Med*. 2016;35:534-552.
10. Taves DR. Minimization: a new method of assigning patients to treatment and control groups. *Clin Pharmacol Ther*. 1974;15:443-453.
11. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*. 1975;31:103-115.
12. Scott NW, McPherson GC, Ramsay CR, Campbell MK. The method of minimization for allocation to clinical trials: a review. *Control Clin Trials*. 2002;23:662-674.
13. Treasure T, MacRae KD. Minimisation: the platinum standard for trials? Randomisation doesn't guarantee similarity of groups; minimisation does. *BMJ*. 1998;317:362-363.
14. Yarnold PR, Soltysik RC. *Maximizing Predictive Accuracy*. Chicago, IL: ODA Books, 2016. https://doi.org/10.13140/RG.2.1.1368.3286
15. Yarnold PR, Soltysik RC, Bennett CL. Predicting in-hospital mortality of patients with AIDS-related *Pneumocystis carinii* pneumonia: an example of hierarchically optimal classification tree analysis. *Stat Med*. 1997;16:1451-1463.
16. Linden A, Butterworth SW. A comprehensive hospital-based intervention to reduce readmissions for chronically ill patients: a randomized controlled trial. *American Journal of Managed Care*. 2014;20:783-792.
17. Linden A, Butterworth S, Roberts N. Disease management interventions II: what else is in the black box? *Dis Manag*. 2006;9:73-85.
18. Linden A, Butterworth SW, Prochaska JO. Motivational interviewing-based health coaching as a chronic care intervention. *J Eval Clin Pract*. 2010;16:166-174.
19. Biuso TJ, Butterworth S, Linden A. Targeting prediabetes with lifestyle, clinical and behavioral management interventions. *Dis Manag*. 2007;10:6-15.
20. Yarnold PR, Soltysik RC. Theoretical distributions of optima for univariate discrimination of random data. *Decision Sciences*. 1991;22:739-752.
21. Yarnold PR, Soltysik RC. *Optimal Data Analysis: Guidebook with Software for Windows*. Washington, D.C.: APA Books; 2005.
22. Linden A, Adams J, Roberts N. The generalizability of disease management program results: getting from here to there. *Manag Care Interface*. 2004;17:38-45.
23. Yarnold PR. Discriminating geriatric and non-geriatric patients using functional status information: an example of classification tree analysis via UniODA. *Educ Psychol Meas*. 1996;56:656-667.
24. Linden A, Yarnold PR. Using data mining techniques to characterize participation in observational studies. *J Eval Clin Pract*. 2016;22:839-847.
25. Ryan P. Random allocation of treatments in blocks. *Stata Technical Bulletin*. 1998;41:43-46.
26. Ryan P. RCT_MINIM: allocation of treatments to balance prognostic factors in controlled trials. Statistical Software Components s457029, Boston College Department of Economics, 2017. Available at https://ideas.repec.org/c/boc/bocode/s457029.html [Accessed on 12 May 2017].
27. Linden A, Yarnold PR. Using machine learning to assess covariate balance in matching studies. *J Eval Clin Pract*. 2016;22:848-854.
28. Linden A, Yarnold PR. Using machine learning to identify structural breaks in single-group interrupted time series designs. *J Eval Clin Pract*. 2016;22:855-859.
29. Linden A, Yarnold PR, Nallomothu BK. Using machine learning to model dose-response relationships. *J Eval Clin Pract*. 2016;22:860-867.

30. Linden A, Yarnold PR. Combining machine learning and matching techniques to improve causal inference in program evaluation. *J Eval Clin Pract*. 2016;22:868-874.

31. Linden A, Yarnold PR. Combining machine learning and propensity score weighting to estimate causal effects in multivalued treatments. *J Eval Clin Pract*. 2016;22:875-885.

32. Linden A, Yarnold PR. Using classification tree analysis to generate propensity score weights. *J Eval Clin Pract*. https://doi.org/10.1111/jep.12744

33. Linden A, Yarnold PR. Using classification tree analysis to model time-to-event (survival) data. *J Eval Clin Pract*. https://doi.org/10.1111/jep.12779

How to cite this article: Linden A, Yarnold PR. Minimizing imbalances on patient characteristics between treatment groups in randomized trials using classification tree analysis. *J Eval Clin Pract*. 2017;23:1309–1315. https://doi.org/10.1111/jep.12792